JU9406V1

SDPS Performance Requirements Interpretation

Working Paper

June 1994

Prepared Under Contract NAS5-60000

RESPONSIBLE ENGINEER

George Percivall, Systems Engineer — Date
EOSDIS Core System Project

SUBMITTED BY

P. G. O'Neill /s/ — 6-24-94

Pete O'Neill, SI&P Office Manager — Date
EOSDIS Core System Project

Hughes Applied Information Systems
Landover, Maryland

This page intentionally left blank.

Contents

This page intentionally left blank.

1. Introduction

1.1 Purpose

This note summarizes the ECS program's interpretation of the performance requirements for use in the analysis of the SDPS design. The interpretation described after each requirement does not imply a modification to the requirement; the baseline requirements apply in all cases.

1.2 Organization

This paper is organized as follows:

Section 1 provides information about the white papers purpose and who to contact for more information.

Section 2 provides interpretations of F&PRS requirements applicable to SDPS. Section 2 is organized into four sections which are the sections of the F&PRS where SDPS applicable requirements are found: Overall Requirements (F&PRS Section 5.3), SDPS Requirements (F&PRS Section 7.4), PGS Performance (F&PRS Section 7.5.1.5.1.4), DADS Performance (F&PRS Section 7.5.1.5.2.4), and IMS Performance (F&PRS Section 7.5.2.4).

An appendix contains the Product Set Data Volumes and Processing Loads as used for SDR analyses.

1.3 Review and Approval

This White Paper is an informal document approved at the Office Manager level. It does not require formal Government review or approval; however, it is submitted with the intent that review and comments will be forthcoming.

The ideas expressed in this White Paper were prepared leading up to the System Design Review and may change with discussions at the review. Although there is presently no plan to migrate these requirements to CDRLs, the interpretations contained herein may become the basis of configuration change requests to the ECS F&PRS document.

Questions regarding technical information contained within this paper should be addressed to the following ECS contacts:

- Mark Elkington, 1616A McCormick Drive, Landover, MD 20785.
  Email: marke@eos.hitc.com Voice: (301) 925-0379 FAX: (301) 925-0327.
- George Percivall, 1616A McCormick Drive, Landover, MD 20785.
  Email: gperciva@eos.hitc.com Voice: (301) 925-0368 FAX: (301) 925-0327.

Questions concerning distribution or control of this document should be addressed to:

Data Management Office
The ECS Project Office
Hughes Applied Information Systems
1616A McCormick Dr.
Landover, MD 20785

2. SDPS Performance Requirements Interpretation

The following sections list requirements from the ECS Functional and Performance Requirements Specification (F&PRS) which apply to SDPS followed by an ECS program interpretation and a relaxation impact.

For each requirement an estimate of the possible impact of relaxing the requirement beyond our interpretation is given in terms of small (<$100K), medium ($100-500K) and large (>$500K).

Throughout the discussion, reference is made to the 'modeling' baseline which differs from the contractual baseline. The modeling baseline is based on the SPSO information delivered in February 1994 and subsequently updated with new information on MISR processing. This baseline is attached as an Appendix. It represents approximately a 30 times increase in processing capacity and 3.5 times increase in storage capacity over the contractual baseline and is referred to as the IWG*1 processing load.

2.1. Overall Requirements (F&PRS Section 5.3)

EOSD1010 ECS shall support daily data volume, processing load, storage volume, instrument support, and data traffic as derivable from and specified in Appendix C and D.

The values of the parameters listed in this requirement have been modified in accordance with the IWG*1 processing load

Relaxation Impact: High (but incurs serious impact on science)

EOSD1040 ECS shall provide sufficient capacity to permit the reprocessing of all EOS science data at twice the incoming data rate at a minimum, concurrently with processing of new data.

This requirement mirrors requirement PGS-1300 below. EDS1040 also affects Ingest whereas PGS-1300 only affects Data Processing (appendix C).

Relaxation Impact: As in PGS-1300, High (but incurs serious impact on science)

EOSD1050 ECS shall generate and make available to the users Level 1 Standard Products within 24 hours after the availability to ECS of all necessary input data sets.

EOSD1060 ECS shall generate and make available to the users Level 2 Standard Products within 24 hours after the availability to ECS of all necessary Level 1 and other input data sets.

EOSD1070 ECS shall generate and make available to the users Level 3 Standard Products within 24 hours after the availability to ECS of all necessary Level 2 and other input data sets.

EOSD1080 ECS shall generate and make available to the users Level 4 Standard Products within one week after the availability to ECS of all necessary Level 3 and other input data sets.

Requirements EOSD1050, EOSD1060, EOSD1070, and EOSD1080 are being used as basic assumptions in the push-side modeling of the system.

Relaxation Impact: Small, since we must process 1 days worth of data in a single day just to keep up. There could be a reduction in GFE inter-DAAC network cost if the dependent data products could be shipped via media.

2.2. SDPS Requirements (F&PRS Section 7.4)

SDPS0120 The SDPS shall be capable of operating in a 24-hour a day, 7-day a week mode.

This requirement needs to be considered in the context of the M&O staffing to be provided. If the system is does not become "hands-off" then this is probably the lowest cost option from the perspective of system development and maximization of throughput. If the system is to be run "hand-off" - i.e. M&O does not provide 24hr/7day operations then either the cost of development will rise to provide a system that can be run unattended or more hardware will be required to allow a higher throughput when it is attended.

Relaxation Impact: None - this is the best solution as long as M&O is available.

SDPS0150 The SDPS shall have the capability of generating quick-look products within 6 hours of receipt of the necessary input data for 10% of the EOS instrument data requiring processing capacity of no more than 10% of the processing requirement for the equivalent standard product.

The impact of this is proportional to the volume of data to be processed. I take this to read that we have to do 10% of the data and the processing only requires one/tenth of the standard processing. This means that the QL processing is using up an equivalent of 1% of the standard processing resource. This seems to be within the error bar of the estimation of the processing requirements. However this is another 'special' requirement that has to be tracked, and verified and of course the estimate of the processing and required throughput might be inaccurate. QL processing should not be done at a DAAC, it should be done on specialized QL processors at EDOS.

Relaxation Impact: High

SDPS0160 The SDPS shall have the capability of generating quick-look products within 1 hour of receipt of the necessary input data for 1% of the EOS instrument data requiring processing capacity of no more than 1% of the processing requirement for the equivalent standard product.

Comments: Given the low volume and relatively low processing required, this requirement is not too onerous as long as it specifies "an average of 1 hour", otherwise special measures to guarantee the RMA requirements will be required.

Relaxation Impact: Medium

2.3. PGS Performance (F&PRS Section 7.5.1.5.1.4)

PGS-1310 The processing capacity necessary to process all EOS science data for which each PGS is responsible shall be based on the data volumes and at-launch instrument processing load requirements (MFLOPS) assigned to each DAAC as well as the 20% yearly product growth as specified in Appendix C.

This requirement is being used as a basic assumption in the push-side modeling of the system. The capacities are derived from the modeling baseline.

Relaxation Impact: High (but incurs serious impact on science)

PGS-1315 Each PGS shall have the capacity to support I/O to temporary and intermediate storage or multiple passes over input products as required by individual science algorithms.

Functionally this requirement is met. The performance modeling assumption for multiple passes over the input products is only to read the data once. From initial discussions with instrument teams on their plans for algorithms this appears to be a reasonable estimate.

Relaxation Impact: Small, because we are already assuming only reading the data once. Although, due to the nature of the algorithms driving this requirement, relaxing this requirement can have serious impacts on science.

PGS-1300 Each PGS shall provide a processing capacity four times the size necessary to process all EOS science data for which it is responsible, including interdisciplinary investigator processing. It shall be possible to effectively utilize the entire reprocessing capacity at each site on computers with similar architectural design (e.g., parallel processors), for a single algorithm or any mix of algorithms normally run at that site. The four times processing capacity accounts for:

a. 1 times to allow for normal processing demands
b. 2 times to allow for reprocessing demands
c. 1 times to allow for algorithm integration and test demands, production of prototype products, ad hoc processing for "dynamic browse" or new search techniques developed by science users, quick-look processing, and additional loads due to spacecraft overlap.

This requirement is being used as a basic assumption in the push-side modeling of the system. The processing capacities are derived from the modeling baseline. It is not assumed that this processing capacity will be provided by launch. The general schedule is as follows:

0.5x	launch-2 years (to support early algorithm I&T)
1.5x	launch-17 months (to support later I&T and launch processing)
1x	launch +12 months (1st reprocessing string)
1x	launch +24 months (2nd reprocessing string)

Relaxation Impact: High (but incurs serious impact on science)

PGS-1301 The effective CPU processing rates used for sizing purposes in PGS-1300 shall not be greater than 25% of peak-related CPU capacity.

This requirement is being used as a basic assumption in the push-side modeling of the system. The processing rates are derived from the modeling baseline. However, it was discovered that the information provided in the SPSO does, in some cases, also include an allowance for processor efficiency. The extent of this double accounting was assessed directly with the instrument teams and corresponding corrections have been made to the defined processing rates in the SDR push-side modeling baseline.

Relaxation Impact: High (but the 25% efficiency is a sensible estimate)

PGS-1270 The PGS design and implementation shall have the flexibility to accommodate PGS expansion up to a factor of 3 in the processing capacity with no changes to the processing design, and up to a factor of 10 without major changes to the processing design. Such expansion in capacity or capability shall be transparent to existing algorithms or product specifications. This requirement shall apply to the system at all phases of contract performance, including the final system which accommodates the product growth specified in Appendix C, as well as the at-launch system.

The modeling baseline already includes a 30 times processing and 3.5 times storage increase. This is referenced as IWG*1 in the SDS. To indicate scalability beyond that level modeling is also being performed to assess 8 times this processing level (i.e., 240 times the contractual baseline) and 2 times storage (i.e., 7 times the contractual baseline; this is referenced as IWG*8. This more than adequately covers this requirement.

Relaxation Impact: Small (this is really only a design issue and therefore the cost of the increased processing is not included).

2.4. DADS Performance (F&PRS Section 7.5.1.5.2.4)

DADS2770 Upon receipt and approval of a request, the designated DADS shall make stored data products available for delivery to the requester within 24 hours for data distributed on physical media.

Using automated media, the time it will take to write data products to physical media will be determined by intrinsic hard media I/O rates and the number of available hard media drives. The slower the media, the more drives we will need (the type of hard media to be used has yet to be determined).

Also, the time it takes to recover from a large backlog of requests will be determined by the number of drives we have, over and above what we need to handle average processing loads. In other words, if we are willing to have a number of hard media drives idle most of the time, we can recover from a backlog more quickly. If we must be able to handle twice the average daily volume in a single day, then we need twice as many drives.

Without specifying a maximum daily distribution volume, the current requirement is not testable.

We should relax this requirement in two ways:

- deliver within an average of 24 hours, 48 hours, etc.
- make the requirement apply to a maximum daily distribution volume (this makes the requirement testable). See DADS3110 as this requirement must be changed also.

Relaxation impact: Large

DADS2778 Each DADS shall be capable of receiving and archiving three days' worth of data (see Appendix C) in any given day.

This requirement is being addressed in the sizing of the Ingest Subsystem and the Data Server Subsystem (e.g., number of drives to support triple the normal archive I/O rate).

Relaxation Impact: High (but would probably require increased buffering capability at EDOS/SDPF/Landsat to compensate) or we could not keep up with the reprocessing load.

DADS2780 Each DADS shall be capable of ingesting data at the maximum output bandwidth of the EDOS.

This requirement is being addressed in the sizing of the Ingest Subsystem.

This is related to a larger issue of having predictable deliveries of data from EDOS. If EDOS can transmit data more or less continuously to ECS (at a reduced I/O bandwidth), then ECS does not have to build as large a buffering capability. This issue should be approached by looking at EDOS and ECS as a single system. Then we could determine where economies could be achieved.

Relaxation Impact: High (but would probably require increased buffering capability at EDOS to compensate)

DADS2900 Each DADS shall provide archival capacity for current volume requirements plus one year. Volume requirements are specified in Appendix C.

This is assumed to imply sizing for the IWG*1 and IWG*8 cases through to the end of contract plus one year.

Processing vs. storage study indicates that with large number of processing dependencies the increase in processor capacity necessary to generate products on-demand would negate any archival savings. This should be re-addressed if the processing dependencies are significantly reduced.

We don't want to buy more archive hardware than we absolutely need to meet current requirements (plus a reasonable amount to respond to significant changes in storage volumes). This is because the technology is changing so much. At issue here is hundreds of terabytes of unused capacity over the life of the

contract. We need flexibility to choose how much extra storage capacity we have.

Relaxation Impact: Large

DADS3000 To support archival data integrity, the bit error rate after correction shall be less than 1 in 10^{12}.

Currently the vendors are signing up to 10^{-12} BER; however, drives providing a 10^{-14} BER are expected to be available. We will select the best technology available at the time of our procurements. The impact of compression techniques we are considering may lower the 10^{-14} back down while still meeting the 10^{-12} BER.

The requirement as now written might be interpreted as an end-to-end requirement. We interpret this (as it was originally written) to be a component requirement.

Relaxation Impact: We are not interested in relaxing this requirement, since low BER reduces operational costs.

DADS3010(1) Archival and backup media at each DADS shall have a manufacture-rated shelf life of at least 10 years when stored in a controlled environment.

This is a typical requirement for archive systems. The shelf-life of the tape technologies that we anticipate using is consistent with this requirement.

Relaxation Impact: None. We are not interested in relaxing this requirement, since long shelf life results in low operational costs due to less frequent media refresh.

DADS3090 Each DADS shall be capable of 200% expansion in throughput and archive capacity without architecture or design change. This expansion capacity shall apply to the total of the at-launch requirement plus the yearly growth requirement specified in Appendix C.

This requirement in terms of archive capacity is more than covered by the IWG*1 and IWG*8 modeling described for PGS-1270. Since the DADS as an element no longer exists in the architecture the throughput is less easy to verify. The throughput for the Data Server which encompasses many of the original DADS requirements has to support throughput from both the push and pull side of the system. The push-side will include the increases implied by the modeling. The pull-side is being derived from a static parametric model which indicates a 3 to 8 times increase over the contractual baseline. The sensitivity of the model to 200% increase on the pull-side will be included in the design analysis.

Relaxation Impact: Small (this is really only a design issue and therefore the cost of the increased archiving is not included).

DADS3100 Each DADS shall be capable of transmitting data over communications network in support of data production requests at the data rate specified in Appendix C

and in support of data distribution requests at a rate equivalent to daily product volume (L1-L4).

The segment design is being analyzed against the above requirement (i.e. IWG*1) and at a level of 4 times larger than that to provide a range of analyses.

Related issue: The recent emphasis on "user pull" has increased the amount of staging disk that we will require. The length of time that products remain in the "user pull" area will have major cost impact.

Relaxation Impact: High (but the purpose of EOSDIS is maximize the effective use of the data archived within it. Should not relax this requirement unless it can be demonstrated that this level of data delivery will swamp the user community leading to ineffective utilization).

DADS3110 Each DADS shall be capable of distributing data via physical media at a rate equivalent to the rate data are ingested (L0) at that DADS.

See comments for DADS2770.

Also, the customer thinks that the real requirement (as specified in the original F&PRS) is to distribute volumes equivalent to L0-L4.

Relaxation Impact: Large

DADS3115 Each DADS shall be capable of making quick-look products available for distribution within 1 minute of receipt from the PGS.

The impact of this requirement depends on the volume (we will have to distribute from SDPS staging). The time period should be stated in terms of average time for availability.

Relaxation Impact:

DADS3120 Each DADS shall distribute product QA data produced at the collocated PGS within 1 hour from the time it is ready.

Is this requirement still valid, since the user's want to "pull" the data?

We assume that the data server is responsible for ingesting the QA data and notifying the SCF of its availability (by means of a subscription).

There is a definite cost to this requirement since it means that we will have to engineer the handling of this data differently than other data. At the very least we will have to give this data higher archiving priority.

The impact of this requirement depends on the volume (we will have to distribute from SDPS staging). The time period should be stated in terms of average time for availability.

Relaxation Impact: Small to Medium

DADS3125 Each DADS shall make archive data, associated with a pre-defined ECS standard format, that is requested for communications network delivery available to the

network in that ECS standard format within an average of 2 minutes after the receipt of a request for that data.

Our interpretation of the two-minute requirement assumes that "archive data" means the first granule for a given request, not all of the data associated with a given request.

Depending on the request, and the load on the system, a given request will not necessarily be handled in a serial fashion; we will handle multiple requests from a single user in parallel, just as we would handle multiple requests from multiple users. Therefore, the second request in a series of requests should take no longer than two additional minutes to complete.

The amount of hardware required to meet this performance requirement depends on 1) the number of requests per day, 2) the total volume of data requested, and 3) the variance associated with both the request arrival times and the volume associated with each request. The variances are the key issue. If there are no variances, then the amount of hardware required to keep up with volume requirements (i.e. do a days work, in one day) will provide performance close to the theoretical best performance of the drives and robotics hardware (i.e. two minutes). Expected variances in arrival times and volumes may necessitate the use of additional hardware in order to provide the required level of performance. Currently we do not know what the variance will be well enough to give an accurate assessment of the cost impact of the two-minute retrieval requirement.

Requiring two-minute performance 100% of the time will be significantly more expensive than requiring that performance 80% of the time (for example).

Relaxation Impact: High

DADS3126 Each DADS shall make archive data, associated with a pre-defined ECS standard format, that is requested for communications network delivery available to the network in a different ECS standard format within an average of 5 minutes after the request for that data.

This requirement is interpreted much as DADS3125 except that some additional manipulation is required on the product before it is sent. The 5 minutes average is for processing done within the Data Server, and does not include any processes which the DAAC may deem are more efficiently done as part of the planned processing. If a user requests a number of items, the average time for the first of the items to start being delivered will be 5 minutes. There should not be a delay of more than 5 minutes between the completion of one item and the start of the next.

The time is determined by the manipulation algorithm. If we can process at I/O rates then we can achieve two-minutes, assuming that there is no queuing time associated with the processor. If the algorithms vary widely in their processing times, the amount of hardware required to achieve five minutes could be great.

Relaxation Impact: High (the ability to smooth processing loads over periods longer than a few minutes will reduce the processing capacity and storage requirements for a data server. We cannot estimate the impact of relaxing this requirement without knowing more about the algorithms.)

DADS3135 The DADS shall have the capability to support the transaction rate as specified in Table 3 (F&PRS Table 7-4).

Comments in Table 3 (F&PRS Table 7-4) apply to data server, which replaces the DADS.

Relaxation Impact: See Table 3 (F&PRS Table 7-4).

2.5. IMS Performance (F&PRS Section 7.5.2.4)

The inventory size is based on the number of data sets and granules and size of granules given in Appendix C. For the purposes of IMS performance estimation, guides should be sized as 5000 documents of 50 pages each (average). A directory of at least 10,000 data set entries and an average record size of 2500 bytes per entry should be used in performance estimation. The user accounting system will be sized to track information on 100,000 users, and in request status checking, it should be assumed that there are 2000 outstanding requests for data.

The numbers based on user modeling will be used in place of these numbers in sizing the system design. Not all of these numbers are possible to directly map to the system design. The directory capacity, guide size, number of users and outstanding requests are assumed to be across the entire system.

Relaxation Impact: small-medium

The usage load for determining system performance is estimated from the number of orders and queries placed at data centers today and adding in a percentage for growth. Given this, there will be an average of 400 IMS queries per hour (including user authorization checks, inventory, directory, and guide queries, and request status queries). In addition, the IMS must support the browse, document search, and ordering activities. For performance criteria, a load of 100 concurrent IMS sessions distributed across the DAACs will be assumed. These concurrent sessions will be distributed across the 8 DAACs in proportion to the projected use and activity at the various DAAC sites. In testing performance, level of activity (i.e., number of IMS operations per hour) should be at least that described in Table 3 (F&PRS Table 7-4). Note that the number of operations in any category will be distributed among the specific operations indicated.

The user modeling activity indicates some values which are different from the above assumptions. We propose to use a usage load characterized from an analysis of the ECS user modeling effort (See Tables 1 and 2).

The design does not recognize an IMS. The functions of the original IMS are now spread between the client, interoperability, data management and data server subsystems. The implication of the above performance characteristics mainly relate to the Data Management and Data Server subsystems. The values that relate to those listed above are:

A parametric model is being used to map the usage load characteristics onto the system design.

Relaxation Impact: High (particularly if the full non-science user load is considered)

Table 1. Daily User Accesses (Minimum)

Daily User Accesses (MIN)	EOS	Other Science	non-Science	Total
ASF	0	0	0	0
EDC	270	486	772	1,527
GSFC	304	531	481	1,317
JPL	0	0	0	0
LaRC	463	442	115	1,020
MSFC	17	21	308	347
NSIDC	45	61	160	267
ORNL	0	0	0	0
Total	1,100.0	1,541.6	1,836.0	4,478

Table 2. Daily User Accesses (Maximum)

Daily User Accesses (MAX)	EOS	Other Science	non-Science	Total
ASF	0	0	0	0
EDC	988	1,706	5,435	8,130
GSFC	1,023	1,867	3,412	6,302
JPL	0	0	0	0
LaRC	1,423	1,551	850	3,825
MSFC	59	75	2,152	2,286
NSIDC	176	215	1,150	1,542
ORNL	0	0	0	0
Total	3,670.0	5,414.4	13,000.0	22,084

IMS-1780 The IMS shall respond to each user session operation within the time period specified in Table 3 (F&PRS Table 7-4) with the specified rate of IMS operations.

See Table 3 (F&PRS Table 7-4).

IMS-1785 The IMS performance specified in Table 3 (F&PRS Table 7-4) shall be maintained during other IMS operational activities such as database updates from the DADS.

See Table 3 (F&PRS Table 7-4).

IMS-1790 The IMS shall provide, based upon the data model defined in Appendix C, sufficient storage for, at a minimum:

a. Directory metadata
b. Guide (documentation/reference material) metadata
c. Inventory metadata
d. System space, LSM data, and data base system overhead
e. Metadata staging area
f. Spacecraft housekeeping and ancillary data metadata
g. Science processing library software metadata
h. Summary data statistics.

It is assumed that the purpose of this requirement is to identify sufficient 'disk' space to allow reasonable access to these items. We believe that other data will need to be maintained with high accessibility.

We will also need disk space for browse data, internal file/data set directories, and archive hardware housekeeping data.

Relaxation Impact: Small (to meet other performance requirements we will have to be able to store a lot more than this)

IMS-1800 The IMS design and implementation shall have the flexibility to accommodate 100% expansion in processing and storage capacity without major changes to the IMS hardware and software design. This expansion capacity shall apply to the total at-launch requirement plus the yearly product growth requirement specified in Appendix C.

There is no longer an IMS in the ECS design. We interpret this to mean a 100% growth in the size of the higher level data types (which are commonly used as metadata), the number of transactions and the manipulations performed. Some of this growth is already represented in the figures derived from the user modeling. It is important to recognize that the growth could be significantly more than this if the system is a success and unlimited access is permitted.

Relaxation Impact: Small-Medium (this is really only a design issue and therefore the cost of the increased processing and storage capacity is not included. There is some increased complexity of design).

Table 3 . IMS User Load and Concurrent Session Characteristics (F&PRS Table 7-4) – (Part 1 of 4)

Session Category	Number of IMS Operations per Hour	Specific Operation	Response Time Requirement*	Response Time Design Goal*
Log-on and Authorization	100	Account confirmation and authorization	13 sec	6 sec
"log-on" is not a useful parameter for the new design. The design considers user accesses whether they be interactive or machine-to-machine. User access will be accompanied by user authorization Our assumption for the number of interactive system accesses for sizing purposes is 4500 per day which equates to an approximate peak (/8*1.5) of 800 accesses per hour. The number of machine-to-machine system accesses is TBD. The response time requirement is interpreted to mean an average and not be a limit. The user authorization process may take longer than this for a new user. Relaxation Impact: Small (and probably not realistic)				
Directory Search	80	Search by single keyword attribute	8 sec	2 sec
		Search by multiple keyword and time or space range check	13 sec	7 sec
The directory search will mainly be replaced by queries on the advertising service and access to directory data. Some of the searching within the advertising service may be better achieved through user driven hyperlinks. This requirement is interpreted to mean the number of interactions on the advertising service for the number of operations per hour and the timing are based on directory data searching. The data load from the user characterization is significantly larger than the proposed performance requirement. Assuming 13,000 science users, and assuming that they browse the advertisements once or twice a week (using 1.5/week avg.), this generates 4,000 sessions a day for the family of servers. A session could easily consist of 10 or more interactions (using 15 queries in the following), so this could generate around 60,000 queries a day. Spreading this over 8 hours and applying a 1.5 multiplier (to account for peak usage) would result in about 3 queries/second - or around 10,000 per hour, collectively for all servers. Small sites may have very small loads. Relaxation Impact: Medium (but only regarding the response times and these seem to be a reasonable response)				

Table 3 . IMS User Load and Concurrent Session Characteristics (F&PRS Table 7-4) – (Part 2 of 4)

Session Category	Number of IMS Operations per Hour	Specific Operation	Response Time Requirement*	Response Time Design Goal*
Guide Search	40	Search for document by keyword	8 sec	5 sec
This requirement is interpreted to mean querying on indexed keywords within the guide documents. Other considerations would be text searching (see Document Search) and hyperlinking (within a local site and cross-sites). The text searching requirements within the Guide will be as for the Document Search. A working assumption for the hyperlinking is that initiation of links (i.e. access to linked item) should not take more than 5 seconds. The usage load derived from considering the user characteristics is TBD. Relaxation Impact: Small				
Inventory Search	120	Search one instrument by multiple keyword attribute w/ time or space range check (one DAAC)	8 sec	2 sec
		Search multiple instruments by multiple keyword attributed w/ time or space range check (one DAAC)	18 sec	7 sec
		Multiple DAAC inventory search by keyword attributes and time and/or space range check	58 sec	11 sec
The results from the user model indicate that the majority of users expect to receive inventory queries within 60-120 seconds. The distribution of expected response did not reflect the complexity of the query as expressed in the categories above. The number of inventory queries per hour is estimated to be TBD. The first category of query is interpreted to be a Data Server only query; the second a LIM + Data Server, and the third would also involve a DIM. More detailed sensitivity analysis using detailed inventory characteristics and representative queries to investigate the difficulty of achieving the above performance. Suggest the representative queries are derived as s test set and agreed with users and appropriate performance targets (by release) are agreed. Relaxation Impact: High (the above response times do not match user expectations as reported in the user scenarios. May require some undesirable simplifications of the query process to meet the second two categories. The response times also do not reflect the size of the various inventories that will exist within ECS)				

Table 3 . IMS User Load and Concurrent Session Characteristics (F&PRS Table 7-4) – (Part 3 of 4)

Session Category	Number of IMS Operations per Hour	Specific Operation	Response Time Requirement*	Response Time Design Goal*
Status Check (account or request)	60	Status of pending order or Data Acquisition Request	13 sec	10 sec
		Account status retrieval	13 sec	6 sec
Interpretation is clear. Relaxation Impact: Small				
Browse *(for data selection)*	50	Retrieve and begin to display standard pre-computed browse product	58 sec	
This is likely to be an expensive performance requirement to fulfill and also a difficult requirement to interpret due to the very different characteristics of browse data - some (e.g., CERES) is generated on demand; others (e.g., LIS) maintain browse data within a product, etc. The number of requests for browse data estimated from the user model is TBD. Suggest that the performance requirement is interpreted as follows: • 'Delivery' for browse products which are separate physical items and are from data received within three months should be within 58 seconds. An equivalent volume of browse data per site should also be accessible with this performance for data that a specific DAAC determines is important. • 'Delivery' of browse data which are separate physical items but not from the previous category (i.e. <3 months old or DAAC-interesting) will have the same performance as any other data item extraction (see DADS3125) • 'Delivery' of browse data which needs to be generated shall follow the same performance requirements as DADS3126 Relaxation Impact: High				
Document Search	10	Search 1000 document pages by keyword	3 sec	3 sec
The requirement is interpreted to mean searching by indexed keyword. Additional performance characteristics are required for non-indexed text string searching and Boolean searching. Relaxation Impact: Small				

Table 3 . IMS User Load and Concurrent Session Characteristics (F&PRS Table 7-4) – (Part 4 of 4)

Session Category	Number of IMS Operations per Hour	Specific Operation	Response Time Requirement*	Response Time Design Goal*
Ordering Services	25	Local DAAC order submission and confirmation	13 sec	12 sec
		Remote DAAC order submission and confirmation	38 sec	30 sec
		Order cost estimate	13 sec	12 sec
Ordering services do not exist in the current design. An order is simply a request which results in data being delivered to the user's site or machine This requirement is assumed to be initial confirmation that order is accepted. Full confirmation (including estimates of processing time and price) should have a longer response time, since this may involve some element of local replanning. Full confirmation within 5 minutes is suggested. The order cost estimate should be replaced by a price estimate. The agreed policy is that user should see consistent price - not the cost of satisfying the request since this removes the transparency of process or store. Relaxation Impact: Medium				
Access Services				
No requirements have been defined for all of the access services that a Data Server will support. These need further consideration based on the user model characterization. Relaxation Impact:				

**(from initiation of query to start of display, exclusive of user environment and network delay)

This page intentionally left blank.

Appendix: Product Set Technical Baseline

This appendix contains the product set technical baseline used for the ECS System Design, June 1994. Two tables are provided; Table A-1 contains the SDR Product Set Data Volumes and Table A-2 contains the SDR Product Set Processing Loads. Both are based on At Launch products (i.e., Post Launch products are not included in the tables).

Table A.-1. SDR Product Set Data Volumes

Level	Platform	DAACs								TOTAL (GB/DAY)
		ASF	EDC	GSFC	JPL	LaRC	MSFC	NSIDC	ORNL	
Level 0	TRMM					0.113	0.065			**0.178**
	AM-1		89.640	66.960		41.326				**197.926**
	COLOR			0.612						**0.612**
	AERO					0.262				**0.262**
	ADEOS II				0.055					**0.055**
	PM-1			82.376		0.221	0.724			**83.321**
	ALT			1.080	0.015					**1.095**
	CHEM			0.540		0.273				**0.813**
	Total		**89.640**	**151.568**	**0.070**	**42.195**	**0.789**			**284.262**
Level 1	TRMM					0.616	0.675			**1.291**
	AM-1		296.000	644.200		200.933				**1141.133**
	COLOR			0.703						**0.703**
	AERO					0.016				**0.016**
	ADEOS II				0.050					**0.050**
	PM-1			675.232		1.232	5.100			**681.564**
	ALT			8.619	0.140					**8.759**
	CHEM			1.130		0.062				**1.192**
	Total		**296.000**	**1329.884**	**0.190**	**202.859**	**5.775**			**1834.708**
Level 2	TRMM					0.239	0.018			**0.257**
	AM-1		617.443	148.904		23.110		2.870		**792.327**
	COLOR			0.461						**0.461**
	AERO					0.002				**0.002**
	ADEOS II				0.452					**0.452**
	PM-1		91.574	149.011		0.478	0.036	2.888		**243.987**
	ALT			0.107	0.014			0.030		**0.151**
	CHEM			0.296		0.002				**0.298**
	Total		**709.017**	**298.779**	**0.466**	**23.831**	**0.054**	**5.788**		**1037.935**
Level 3/4	TRMM					0.043				**0.043**
	AM-1					0.110				**0.110**
	COLOR			0.537						**0.537**
	AERO									
	ADEOS II									
	PM-1					0.086				**0.086**
	ALT				0.001					**0.001**
	CHEM									
	Total			**0.537**	**0.001**	**0.239**				**0.777**
TOTAL (Level 0-4)	TRMM					1.011	0.758			**1.769**
	AM-1		1003.083	860.064		265.479		2.870		**2131.496**
	COLOR			2.313						**2.313**
	AERO					0.280				**0.280**
	ADEOS II				0.557					**0.557**
	PM-1		91.574	906.619		2.017	5.860	2.888		**1008.958**
	ALT			9.806	0.170			0.030		**10.006**
	CHEM			1.966		0.337				**2.303**
	TOTAL		**1094.657**	**1780.768**	**0.727**	**269.124**	**6.618**	**5.788**		**3157.682**

Table A-2. SDR Product Se Procerssing Loads

Level	Platform	DAACs								TOTAL (MFLOPS)
		ASF	EDC	GSFC	JPL	LaRC	MSFC	NSIDC	ORNL	
Level 1	TRMM					1.876	0.883			**2.759**
	AM-1		22.000	2400.000		1468.910				**3890.910**
	COLOR									
	AERO					0.212				**0.212**
	ADEOS II				0.066					**0.066**
	PM-1			2440.240		3.750	0.850			**2444.840**
	ALT			23.150	1.500					**24.650**
	CHEM			15.228		0.219				**15.447**
	Total		**22.000**	**4878.618**	**1.566**	**1474.967**	**1.733**			**6378.884**
Level 2	TRMM					0.188	0.100			**0.288**
	AM-1		366.338	526.714		534.235		2.300		**1429.587**
	COLOR									
	AERO					0.207				**0.207**
	ADEOS II				3.446					**3.446**
	PM-1		234.293	2126.714		0.375	1.200	2.700		**2365.282**
	ALT			1.023	1.500			0.059		**2.582**
	CHEM			2882.800		0.207				**2883.007**
	Total		**600.631**	**5537.251**	**4.946**	**535.212**	**1.300**	**5.059**		**6684.399**
Level 3/4	TRMM					0.094				**0.094**
	AM-1					0.548				**0.548**
	COLOR									
	AERO									
	ADEOS II									
	PM-1					0.188				**0.188**
	ALT				5.000					**5.000**
	CHEM									
	Total				**5.000**	**0.830**				**5.830**
TOTAL (Level 1-4)	TRMM					2.158	0.983			**3.141**
	AM-1		388.338	2926.714		2003.693		2.300		**5321.045**
	COLOR									
	AERO					0.419				**0.419**
	ADEOS II				3.512					**3.512**
	PM-1		234.293	4566.954		4.313	2.050	2.700		**4810.310**
	ALT			24.173	8.000			0.059		**32.232**
	CHEM			2898.028		0.426				**2898.454**
	TOTAL		**622.631**	**10415.869**	**11.512**	**2011.009**	**3.033**	**5.059**		**13069.113**

This page intentionally left blank.

Abbreviations and Acronyms

DADS	Data Archive and Distribution System (ECS)
ECS	EOSDIS Core System
F&PRS	Functional and Performance Requirements Specification
IMS	Information Management System (ECS)
PGS	Product Generation System (ECS)
QL	Quick Look
RMA	reliability, maintainability, availability
SDR	System Design Review
SPSO	Science Processing Support Office